Ranking of Geostatistical Models and Uncertainty Quantification Using Signal Detection Principle (SDP)

Maureen Ani [a,*], Gbenga Oluyemi [a], Andrei Petrovski [b], Sina Rezaei-Gomari [c]
[a] *School of Engineering, Robert Gordon University, Aberdeen, AB10 7QB, UK*
[b] *School of Computing, Robert Gordon University, Aberdeen, AB10 7QB, UK*
[c] *School of Science, Engineering & Design, Teesside University, Middlesbrough, TS1 3BX, UK*

Abstract

The selection of an optimal model from a set of multiple realizations for dynamic reservoir modelling and production forecasts has been a persistent issue for reservoir modelers and decision makers. Current evidence has shown that many presumably good reservoir models which originally matched the true historic data did not always perform well in predicting the future of the reservoir as a result of uncertainties.

In this paper, a new method for optimal model selection using Signal Detection Principle (SDP) approach is presented. In principle, SDP approach models the dissimilarity between various realizations as a cross-function of spatial distance, statistical correlation and the inherent noise level in the model; while the existing methods estimate the dissimilarity between parameter values of different realizations as a function of distance (statistical or spatial distance) or quality factor. SDP approach fills the model divergence gap by way of classifying the information-bearing patterns in a model as signals, and identifying random patterns which bears no information as noise. Thus, it quantifies the mismatch uncertainty and the inherent fudge in every realization as a measure of the model's reliability. The SDP approach has been validated with historical data, and the results show that the reliability factor of the best model has a value of one. Thus, if the strongest reservoir model has a reliability factor of one, then that model will make approximately the closest and best predictions as the true system in every forecast; whereas if the reliability factor of the model is less than or greater than one, its predictions will always disperse from historical data and is unsuitable for reservoir forecasts. Two examples based on real field data are used to demonstrate the application of SDP approach. In these case examples SDP was used to analyse the difference between the reservoir models in a set of multiple realizations. The target variable in this study was STOIIP (Stock Tank Oil Initially in Place) . The results show that the models with a lower reliability factor made similar predictions as the real reservoir data in almost all the forecasts. Similarly, the models with higher reliability factors made dissimilar and more unreliable predictions in most of the forecasts.

Keywords: Reservoir Simulation; Multiple Realizations; Model Ranking; Uncertainty Quantification; Signal-to-Noise Ratio; Dissimilarity Measure.

* Corresponding Author
E-mail address: m.n.ani@rgu.ac.uk

1.0. Introduction

In recent years, the selection of a candidate model from a set of model realizations for the purpose of solving various reservoir modelling problems has attracted a lot of research interests (Shirangi and Durlofsky, 2016). The areas of its application in the oil and gas industry range from history matching (Mtchedlishvili, Voigt and Haefner, 2004; Sahni and Horne, 2004; Shirangi, 2014; Shirangi and Emerick, 2016), well performance (Ballin, Journel, and Aziz, 1992), production forecast (Pratt *et al.*, 1986), reservoir property distribution (Fanchi, 2006), reservoir management (Deutsch and Srinivasan, 1996; Fei *et al.,* 2016), uncertainty analysis (Yadav, Bryant and Srinivasan, 2006; Rahim, Li and Trivedi, 2015; Shirangi and Durlofsky, 2016), well placement optimization (Guerra and Narayanasamy, 2006), to field development planning and decision-making (Hegstad and Saetrom, 2014) amongst others.

Like other type of models, every reservoir model has uncertainties in it. Therefore, after building a reservoir model, multiple realizations of the model are generated in an effort to capture the full range of the uncertainties associated with the model. Conventionally, in a reservoir modelling workflow; it is mostly computationally expensive and futile to carry out flow (dynamic) simulation for all the realizations, thus a reference dynamic model is rather selected from a suite of multiple static models (Shirangi and Durlofsky, 2016). However, the selected model seldom represents the optimal model. Considering that each model has its own independent probability distribution, it is important to consider the individual features of each model to be able to compare one model to another; given that for any pair of model realizations, there is a dissimilarity measure that exists between them.

The focus of this paper is to use a dissimilarity function which is based on the intrinsic features of each model to optimize the selection of the best model from a set of multiple realizations; the one that fits as the most representative model with the least uncertainty for dynamic modelling. Such features include the individual probability distribution of each model, the pattern of inherent information each model is bearing - which could either be a signal or a noise. In the context of this paper, a reference model is a model which has either been standardised with the historical data of the actual reservoir system, or a model which has been calibrated to represent the actual reservoir system as much as possible within reasonable constraints.

1.1. Existing Model Selection Approaches

Largely, the selection of representative of models from a set of multiple realizations works by basically defining a quantifiable ranking measure, which can be used to evaluate each realization (Scheidt and Caers, 2010). The efficiency of any ranking method however depends on the reliability of the method in selecting realizations that closely predict the real reservoir system (Begg and Welsh, 2014), or at least predict the low, mid and high (P_{10}, P_{50} and P_{90}) reserves quartiles as accurately as possible. Several approaches have been applied in the selection of models from a set of multiple realizations (Deutsch and Srinivasan, 1996; Scheidt and Caers, 2010; Shirangi and Durlofsky, 2016; Sahni and Horne, 2004; Mtchedlishvili, Voigt and Haefner, 2004; Fei, Yarus and Chambers, 2016; Dehdari and Deutsch (2012) and Sharifi *et al*., 2014). Model selection by ranking is generally based on the distribution and response of the quantity of interest. While each model selection approach differs from another, the common logic in all is that they all consider a quantifiable difference function which forms the basis for differentiating between the realizations.

In the area of reservoir modelling and optimization, the ranking approaches are broadly classified into two major types - static and dynamic ranking. Static ranking uses the static geological reservoir properties such as

porosity, pore volume, stock-tank oil initially in-place (STOIIP), etc. as a basis for ranking. On the other hand, dynamic ranking uses the properties which are linked to the dynamic and fluid flow behaviour of the reservoir as a criterion for ranking. Such dynamic reservoir properties include relative permeability and connected hydrocarbon volume (CHCV). Even though relative permeability and CHCV are considered flow properties, they do not actually account for the full flow behaviour of the reservoir as much as flow-physics attributes like tracer (Saad, Maroongroge and Kalkomey, 1996), streamline time-of-flight (TOF) and random-walk (McLennan and Deutsch, 2005(b)).

Pratt *et al*. (1986) applied Principal Component Analysis (PCA) in the ranking of well productivity. This approach is based on the principle of orthogonal transformation, where a set of observations are converted into principal components, such that the non-correlated principal component has the largest possible variance, and thus accounts for most of the variability in the data sample. They analysed the variability of coalbed methane gas production profiles of different wells, using the correlation between lineaments and gas production. Due to the fact that lineaments are not always uniformly scaled, they are not reliable static attributes. Therefore, some of their results were inconsistent and showed that the outcomes could be dependent on who mapped the lineaments and how they were mapped, and thus they recommended that the approach should be used with caution. Ballin, Journel, and Aziz, (1992) emerged with the use of a fast simulator to select models for flow simulation. They proposed a way of reducing considerable computing time and analysing uncertainty by using tracer simulation (fast simulator) on multiple realizations, and finally conducting the comprehensive simulation on only the selected realizations. The main idea of model selection by ranking is to find the ideal model with the appropriate set of input parameters that will be able to match the model that has been calibrated with historic data, and thus predicts the profile or future of the true system. However, the parameter set that can fit any calibrated model is not unique (and this is known as the issue of non-uniqueness); this means that different parameter sets from different realizations will be able to match the same calibrated model. In favour of dynamic ranking measures, Deutsch and Srinivasan (1996) pointed out that although no ranking measure had overcome the problem of non-uniqueness, they were hopeful that streamline and random-walk simulation could be promising techniques for improved ranking. In line with this recommendation, Idrobo, Choudhary and Datta-Gupta, 2000; Wang and Kovscek, 2002; and Ates et al., 2003 applied the streamline time-of-flight (TOF) property as a ranking measure based on the idea that TOF reflects the difference in the flow field of different models, and is also directly related to the reservoir sweep efficiency which in turn reflects the interaction between the reservoir heterogeneity and the flow field.

On the other hand, the static ranking measures have been applied by McLennan and Deutsch, 2005(a); Li and Deutsch, 2008; Fenwick and Batycky, 2011; Li, Deutsch and Si, 2012; Sharifi *et al.*, 2014. Although the dynamic ranking measures are simplified, they could undermine the geological complexity of the reservoir. McLennan and Deutsch, (2005(a)) showed that the static measures of local connectivity in a Steam-Assisted Gravity Drainage (SAGD) reservoir were more effective in accurately identifying the realizations that corresponded to the P_{10}, P_{50} and P_{90} respectively, in comparison to other ranking measures. Li and Deutsch, (2008) introduced a static ranking quality factor which ranked realizations according to their connected hydrocarbon volume (CHCV) with consideration to the position of the producing wells, hence considering only the grid cells that are connected to the producing wells. Li, Deutsch and Si, (2012) modified the quality factor to include the geometric average permeability of the grid cells between the cell of interest and the nearest production well. They pointed out that they obtained better results with the modified quality factor, in comparison to ranking based on CHCV. As an improvement to the foregoing, Sharifi *et al.*, (2014) used a fast-marching method (FMM) to quantify the CHCV in a well. Using the correlation between the CHCV and well

production, FMM is able to capture the dynamic connectivity of the reservoir. Using a different perspective, Fenwick and Batycky, (2011) showed an effective framework for analysing reservoir uncertainty using a single measure of the distance between two models. They explained that this distance is established by placing the ensemble of reservoir models in a metric space. This approach is in line with the application of geostatistical or spatial distance measures as applied in this paper.

Generally, for every ranking technique whether it has considered a static or dynamic ranking measure; the common idea is that each has used its ranking measure as the performance indicator for model selection. These performance indicators include distance-based measures (geostatistical, spatial or statistical distance), quality factor measures, and time-based measures. However, with the aim of selecting the optimal model (that is the closest match to the true system) from a set of multiple realizations, bearing in mind the problem of non-uniqueness of parameter set, this paper introduces a new bifocal approach for model selection based on both fitness and reliability - using a divergence and reliability cross function as measure of each model's ability to predict the real reservoir system.

2.0. Methodology and Workflow

The aim of the SDP (Signal Detection Principle) approach is to classify model realizations, quantify the uncertainties in the realizations, rank them based on their performance, and select the optimal realization for dynamic simulation and future forecasts. After ranking the realizations, the most reliable realization is selected using the Reliability Factor (RF).

The workflow starts with building a geological (static) model of the reservoir using available geologic and petrophysical data to establish the reservoir volume. However, considering that no reservoir model is absolutely accurate, it is crucial to carry out uncertainty quantification of the model output. With geostatistical models, a rule of thumb is to match the initial model with the real geologic or reservoir data and use it as the reference model. Then, multiple realizations of the reference model are generated in order to capture the full range of uncertainties in the model. If production forecasts and reservoir performance prediction is the objective, then a dynamic model is built.

For all future forecasts, estimations and dynamic modelling, the best realization is chosen as the base case. Normally, the best or optimal realization is considered to be the realization that matches the real reservoir data or reference model the most. However, since actual experience has shown that the realization that matches the real-life data the most is not always the best for future predictions, SDP functions and equations (discussed in Section 3) are used to rank the realizations by quantifying the underlying noise in each realization using a dissimilarity factor and Signal-to-Noise ratio (SNR). This is then progressed by establishing the reliability of each realization expressed as a Reliability Factor (RF). The RF is a measure of the accuracy of each realization in making realistic predictions of the reservoir; the ranking and selection of the optimal realization is finalised based on the value of their RFs.

Figure 1 is a block diagram of the workflow of the SDP Uncertainty Quantification and Model Ranking Approach.

.

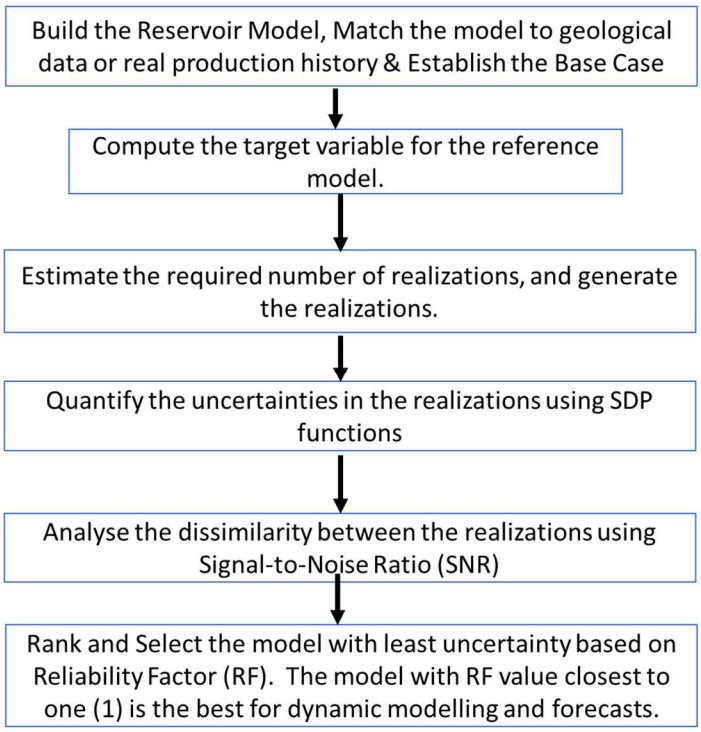

Figure 1: Workflow diagram of the SDP Uncertainty Quantification and Model Ranking Approach

To be able to test the reliability of each of the realizations for both STOIIP estimation and future reservoir performance prediction, the reference model is matched with historic production data. Thus, both STOIIP and production forecast is estimated for each of the realizations.

3.0. Signal Detection Principle (SDP)

SDP is a novel optimization method which differentiates signals from noise in different models by classifying information-bearing patterns as signals, and categorizing random patterns as noises. Given that each model has its own individuality and probability distribution, thus it is possible to compare one realization to another based on their unique attributes. This comparison can be achieved using divergence measures and dissimilarity functions. With the application of the concept of relative divergence, the quality of the information contained in any model can be accurately characterized as either a useful signal or noise. Then, the signal to noise ratio can in turn be used to quantify the inherent uncertainty in each realization or model, and once the models have been characterized, the reliability of each model can be precisely analysed as a function of the model uncertainty and distribution of the variance.

3.1. The Model Divergence Measure

Suppose that A and B are two models in space with varying sets of individual properties θ, and each having its own probability distribution *a* and *b* respectively. Assuming that A is the reference model which has been built to represent the true system (the reservoir) as closely as possible, with a set of multiple realizations B; such that B = B_1, B_2, ..., B_n respectively, where n is the number of elements in the set, and $n > 0$. Given that each model in set B is realized geostatistically, this means that the properties θ in each realization has been spatially distributed across the model grids. Thus, each model assumes that the properties are not just randomly

distributed, but rather statistically correlated in space (Fanchi, 2006). Therefore, in the search for a candidate model for dynamic simulation, the intention will be to achieve models with as little divergence as possible in comparison to the reference model. If A and B have their own probability distributions *a* and *b* which both have their individual outcomes respectively, then the uncertainty (U) associated with each model A and B would be expected to be the sum of their individual probabilities thus:

$$U_{(a,b)} = f(a) + f(b) \quad (1)$$

This additivity in Eq. (1) is valid since the probability distributions of A and B are independent, and if this property holds for all positive integers of a and b, it could be concluded that:

$$f(a) = \log a; f(b) = \log b \quad (2)$$

If each of the various possible outcomes of the models A and B has an individual probability of $\frac{1}{a}$ and $\frac{1}{b}$ respectively, then the dissimilarity measure (di) for each outcome is given as the logarithm of these individual probabilities; expressing this in terms of the probability of model A yields:

$$di = \log\frac{1}{a} \quad (3)$$

In reality, B is an approximated estimate of A. Thus:

$$di = \log\frac{1}{b} \ (\pm approximations) \quad (4)$$

While the base of the logarithm is subjective, the log should be taken to a base that is appropriate to the unit of measurement in use. In this paper, the natural logarithm (ln) is adopted. Whereas in some other cases like in a binary system where there are only two expected outcomes, a log to the base of two might be more appropriate.

To account for the error and divergence from the approximations in B, Equation (4) becomes:

$$di = [\ln\frac{1}{b} + \varepsilon] \quad (5)$$

Then, let the total expected divergence (E) of B from A for the quantity of interest be:

$$E = \sum a * di = \sum a\,[\ln(\frac{1}{b} + \varepsilon)] \quad (6)$$

Thus,
$$E = \sum a\,.\ln\frac{1}{b} + a.\varepsilon \qquad \varepsilon \in [0,1] \quad (7)$$

Given that the term ε was introduced because of the inefficiencies as a result of the uncertainties in B, the last term $(a.\varepsilon)$ in equation (7) is an error term and could be referred to as the approximation uncertainty (R) in this paper. Thus, $= a.\varepsilon$.

Next, it is evident that the first quantity $(\sum a\,.\ln\frac{1}{b})$ in equation (7) is equivalent to entropy. Considering that entropy is a measure of the randomness or uncertainty in random variables and models, therefore Eq. (7) could be expressed in terms of entropy. As a property of entropy, when the probability of the two distributions or

models of interest are the same (in this case when a = b, and $\frac{a}{a} \equiv \frac{a}{b} = 1$), then the relative divergence between the two models is zero. Thus, expressing (7) in terms of entropy by multiplying the divergence function by *'a'* and also dividing by '*a'* yields:

$$E = \sum a \,.\, \ln(\frac{a}{b} . \frac{1}{a}) + R \tag{8}$$

$$E = \sum a \,.\, \ln\left(\frac{a}{b}\right) + \sum a \,.\, \ln\left(\frac{1}{a}\right) + R \tag{9}$$

Therefore, the **total expected divergence (E)** between any realization and the reference model is defined by Eq. (9). Intuitively, Eq. (9) could be fragmented to show that the first term $\sum a \,.\, \ln\left(\frac{a}{b}\right)$ equals relative entropy which is also known as Kullback-Leibler distance ($KL_{a||b}$), the second term $\sum a \,.\, \ln\left(\frac{1}{a}\right)$, equals entropy also known as Shannon's entropy (H), and the last term (R) is the approximation uncertainty.

Consequently, it could be concluded that the total expected divergence (E) of any model realization from the true system is the summation of the relative entropy, the Shannon's entropy and the approximation uncertainty. Thus:

$$E = \sum(KL_{\mathrm{a||b}} + H + R) \tag{10}$$

With equation (10), the actual impact of applying the wrong probability distribution in any reservoir model can be precisely characterized, and the exact impact of the model mismatch uncertainty can also be quantified.

3.2. The Signal-to-Noise Ratio (SNR)

Making inferences and predictions with just a singular and presumably best model is not always plausible (Mtchedlishvill, Voigt and Haetnev, 2004). Thus, it is rather more realistic to quantify the uncertainty and reliability of each fitted model because almost every model has an inherent fudge factor.

Signal-to-Noise (SNR) ratio is a measure of the amount of useful information that is contained in any system; in this case the reservoir model. A popular SNR is the one commonly applied in information and communication systems. However, SNR in information and communications is different from the SNR application in this paper in the sense that the former is a function of the squared signal and noise voltage or power, whereas in the context of this paper, SNR is a function of the approximation uncertainty (R) and the relative variance ($\vartheta^2_{|r|}$) of the random variables.

Statistically, the variance (ϑ^2) is a measure of inconsistency or divergence, and the variance (ϑ^2) of the probability distribution of a set of random variables (of a sample x) is the squared deviation of the distribution from their respective mean ($\bar{x}$). It is a quantity of spread. In most general cases, the variance is expressed as:

$$\vartheta^2 = \frac{\sum(x - \bar{x})^2}{n} \tag{11}$$

The relative variance ($\vartheta^2_{|r|}$) is like the variance probability, and unlike the mean and standard deviation, it proves to be useful for testing a hypothesis within a distribution. It is expressed as:

$$\vartheta_{|r|}^2 = \left(\frac{\vartheta^2 n}{\sum \vartheta^2}\right) \quad (12)$$

The Signal Detection Principle (SDP) method presented in this paper has been used to test the variability associated with the concentration of the elements and attributes that each model has acquired from its input parameters, and this provides the information for analysing the reliability of the model. The total amount of information (i.e. useful and non-useful information) in a given model equals one. Some of the key properties of the relative variance are that its sum in any given set of model distribution is unity (i.e. in any set of multiple model realizations, $\Sigma\vartheta_{|r|}^2 = 1$), and it is a non-negative quantity ($\Sigma\vartheta_{|r|}^2 \geq 0$).

Likewise, the Signal (*S*) in any given model is the amount of useful information (non-variability) in the model, and it is denoted by:

$$S = (1 - \vartheta_{|r|}^2) \qquad S \in [0,1]. \quad (13)$$

Similarly, Noise (*N*) in each model is the amount of non-useful information (variability) in the model, and it is denoted by:

$$N = (1 - S) \qquad N \in [0,1]. \quad (14)$$

The values for both *S* and *N* lies between zero and one in any model, and a model with a very high *S* value (as close as possible to one) is desired. To quantify the reliability of the model, the signal-to-noise (SNR) is determined, and it is simply the ratio of signal to noise:

$$SNR = \frac{S}{N} \quad (15)$$

3.3. Reliability Factor

Once the SNR is established, the Reliability Factor (RF) of the realization is modelled as a function of the approximation uncertainty (R), the SNR and the variance probability (ϑ^2). Reliability factor is given as a fraction (%) and expressed as:

$$RF = R + \left(\frac{SNR}{\sum \vartheta^2}\right) \quad (16)$$

Therefore, the Mutual Divergence ($\overline{D}$) of a realization from its reference model becomes:

$$\overline{D} = (1 - RF) \quad (17)$$

$$\overline{D} = 1 - [R + \left(\frac{SNR}{\sum \vartheta^2}\right)] \quad (18)$$

$\overline{D}$ is expressed as a fraction (%), and is equivalent to the uncertainty of the model's predictability. Note that the positive or negative value in $\overline{D}$ does not matter, it only indicates the direction to which the uncertainty lies with reference to the optimal (selected) model.

4.0. The Models, Case Examples and Result Analysis

The application of the SDP approach for the model ranking and selection is demonstrated in this section. Details of the background of the case study models, their multiple realizations and the obtained results are also given.

4.1. The Reference (Base Case) Reservoir Model

For this study, the target variable is the Stock Tank Oil Initially in-Place (STOIIP) which is completely dependent on the static properties of the reservoir. Hence, a geological model was built for each of the two case reservoirs (Field A and Field B respectively) shown in Section 4.3, to model their rock and fluid properties, as well as carry out the volumetrics. The model was built using the fundamental building blocks for a geological model which are 3D seismic data, fault polygons and well data (well logs, well tops and well deviations). Then, the reservoir properties (oil-water contact depth, gas-oil contact depth, bulk volume and cell volume) were defined in the model. Also, the petrophysical properties of the reservoir (porosity, permeability, facies and water saturation) were modelled; using these information, the volume calculation (STOIIP) was carried out. At this stage, the model was constrained by the reservoir geologic data to ensure that the model agrees with the field properties. The results showed that the STOIIP volumes in the respective models matched the geologic data.

Given that every model is subject to uncertainties which makes its predictions different from real life, the models that were constrained to the geologic data of the two case reservoirs were used as reference models. Essentially, a reference model is the model that qualifies to be taken as the starting point for the reservoir simulation. Thus, in this study, each reference model is the model that has the most likely or best guess values of the input parameters.

4.2. Creating Multiple Realizations to Capture the Full Range of Model Uncertainty

One of the benefits of building a geostatistical reservoir model is to be able to explore and capture the full range of uncertainties in the model, by generating many equally probable cases of the reservoir model from the same reference case. These discrete equiprobable cases are called realizations. Also, both STOIIP and production forecast is estimated for each of the realizations. This is to be able to test the reliability of each of the realizations for both STOIIP estimation and future reservoir performance prediction.

After generating the multiple discrete realizations of the reference model for Field A, then the most realistic realization from the ensemble was selected using the SDP approach. Similarly, the same principle (as in Field A) was also applied in generating the multiple realizations for Field B.

4.2.1. The Rationale for Choosing the Optimum Number of Realizations

One big question that always arises whenever the issue of multiple realizations is discussed is “what number of realizations will be optimal?” The fact is that the number of realizations required to capture the full range of uncertainties within the probability distribution of any model is not indefinite, there is a cut-off for the optimal number of realizations. Choosing the number of realizations without obeying this principle leads to a statistical modelling error that is governed by the principle of underfitting and overfitting. Underfitting occurs when the model does not fit the dataset or its distribution well enough to capture the trend of the data, while overfitting occurs when the model fits the dataset too well that it begins to capture the noise in the dataset as opposed to isolating it.

Generally, the number of realizations considered adequate increases with the spatial variability of the input parameters, as well as the sample size. Ideally, the adequate number of realizations is related to the number of parameters in the model space. In this paper, the optimum number of realizations was determined by carrying out a sensitivity analysis on the Margin of Error (MOE) for different number of realizations. The Margin of Error (MOE) is the boundary or limit of the expected error within the entire range of uncertainty in any model, and the exact number of realizations that corresponds with this margin is the number required to capture the full range of uncertainties in the model. The MOE is also referred to as the *error margin* in this paper and expressed in Equation 19.

$$\text{Margin of Error (MOE)} = \left[\frac{(Confidence\ limit).(Standard\ deviation)}{\sqrt{Number\ of\ Realizations}}\right] \tag{19}$$

Firstly, the sample mean for different number of realizations is obtained. Then depending on the type of distribution for the target variable, the confidence limit (Z-value) is established. The 95% confidence z-value (1.96) was used in this case. The error margin obtained continued to reduce as the number of realizations increased. However, as the number of realizations increased, the error margin reaches a point at which the marginal decrease in the error itself begins to increase. The point at which this effect sets in marks the optimum number of realizations required to completely represent the full range of uncertainties in the reference model. Figures 2 and 3 show the effect of the number of realizations on the error margin for the reservoir models of Fields A and B respectively. Figure 2 reveals that approximately twenty models are required for Field A, thus nineteen realizations were generated for Field A in addition to the reference model, whereas Figure 3 showed that about nine to ten realizations were required for Field B; but then nine more realizations were generated for Field B in addition to its reference model (instead of settling for the upper bound which would have meant generating ten more realizations).

As a rule of thumb, a minimum of 10 realizations is required to achieve a good representation of any model scenario, and to have a good idea of the spread of the uncertainty in the model (Cannon, 2018). While a minimum of 10 realizations are enough to provide a true idea of the spread of the data and the associated uncertainty in a model, it is pertinent to mention that from a full-field geostatistical modelling standpoint, at least 100 realizations will be required. In the case of some giant reservoir models, hundreds to several thousands of realizations could be required to be able to explore the full field uncertainty ranges.

Lastly, it is very significant to mention that the margin of error is a function of the confidence interval and standard deviation within which the mean of the data lies. Just as the mean has the unit of the data, likewise the error margin has the unit of the data (E.g. MMSTB, bbl., etc.)

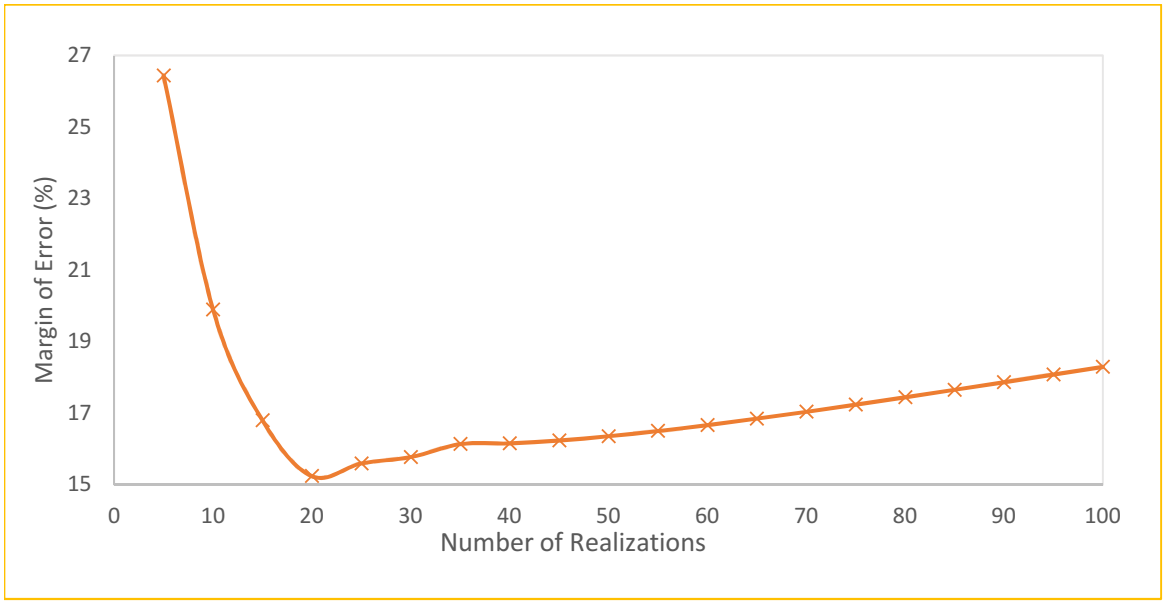

Figure 2: Sensitivity on Realization Count for Field A

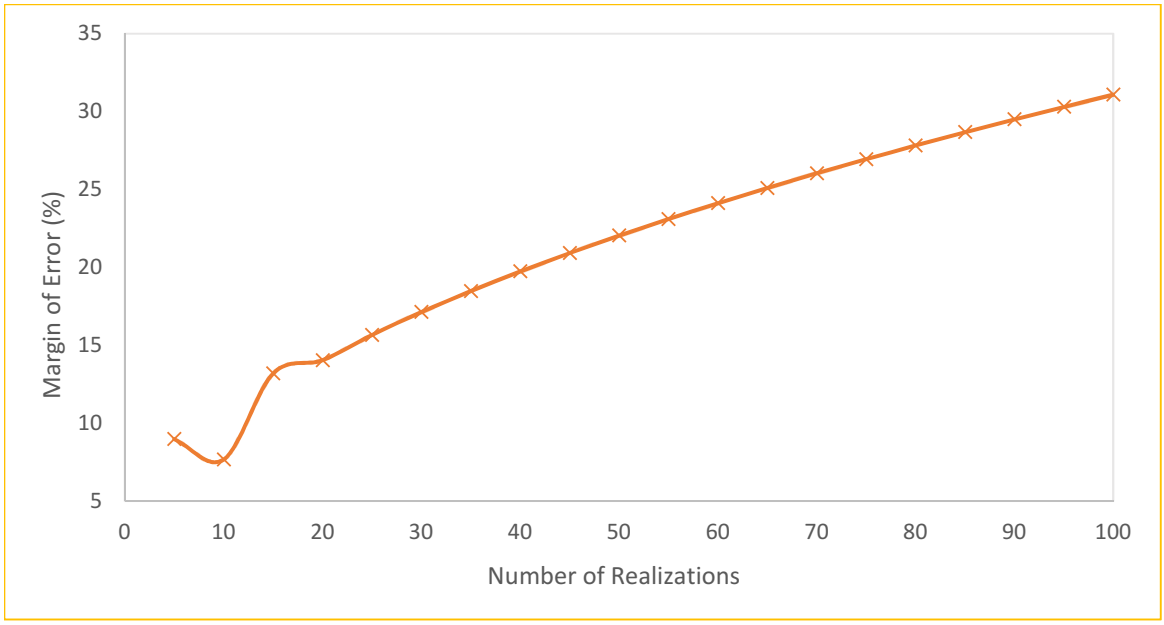

Figure 3: Sensitivity on Realization Count for Field B

4.3. Case Examples

Two examples are presented in this section, in which the ability of the SDP method to select the optimal model from an ensemble of realizations, as well as its ability to analyse the reliability of each model is verified.

4.3.1. Example 1 (Field A)

A set of reservoir data from a North Sea field (Field A) is used. Part of the data set is the historic production data of the field. Field A is an oil and gas offshore field with deltaic sandstones, and it is almost at the end of its life span. For the purpose of this paper, only the oil-bearing reservoir formation of this field is considered. Table 1 shows a brief overview of the reservoir information relevant to this paper

Table 1: Field A model information

Number of Realizations	20 *(i.e. 19 Realizations + Base Case)*
Years of Production (Years)	32
Stock Tank Oil Initially-in-Place (STOIIP), MMSTB	4985
Parameter of interest in model selection	STOIIP

The reservoir model was built and conditioned to fit the actual reservoir production data, and validated as the reference model. Then, nineteen (19) multiple realizations of the reference model were generated. Figure 4 shows the production profile of Field A reservoir, alongside its 19 realizations.

Obviously, it can be seen that most of the realizations in Figure 4 match the reference model very closely. Judging the accuracy of the realizations just on the basis of how closely they fit the reference model without a reliable empirical analysis will not be reasonable, as charts are not able to tell the story behind the spatial correlations of the realizations. Thus, the SDP method was used to analyse the spatial distance, statistical correlation and uncertainties associated with each model. Figure 5 shows the STOIIP distribution of Field A with the reference model (Model 1) and the realizations (Models 2 to 20). A trendline fit to the STOIIP of the reference model shows four realizations (14, 16, 17 and 20) whose predicted STOIIP closely matches the reference model. However, the SDP model shows which of these models is a better candidate.

Using the STOIIP and the probability distribution of the reference model and its realizations as input in the SDP model, the total expected divergence of each model was realized, and the initial ranking was done based on the measure of divergence of each realization from the reference model.

Once the total expected divergence (E) of each realization from the reference model was obtained, the signal-to-noise ratio (SNR) was applied to verify the reliability of each model prior to making final decision. Thus, the amount of useful information (signal) and noise in each realization was analysed (including the reference model, for the benefit of quantifying the model uncertainty). With the SNR, the reliability factor (RF) and associated uncertainties of each model are analysed. Table 3 shows the results for Field A.

Ranking with only the total expected divergence (E) of each realization showed that for Field A; realizations 14, 17, 16 and 20 ranked as the top four respectively on the basis of their expected distance from the true system. However, applying the RF showed that although realization 14 matched the reference model the most, it had more uncertainties than realization 17; and thus realization 17 would rather make a closer and more similar prediction of the reservoir in comparison to 14 and the other realizations.

4.3.2. Example 2 (Field B)

For the purpose of validating the SDP method, it was applied on a different reservoir (Field B) from a completely different geologic setting. Field B is an oil-bearing sandstone reservoir located onshore West Africa in the Niger-Delta basin with a STOIIP of 353 MMSTB, and its reservoir depth is about 7000 ftss. This field is currently in its mid-life and the cumulative production profile is also considered. Table 2 shows a brief overview of the reservoir information, and Figure 6 shows the production profile of Field B, together with its eight realizations.

Table 2: Field B model information

Number of Realizations	10 *(i.e. 9 Realizations + Base Case)*
Years of Production (Years)	18
Stock Tank Oil Initially-in-Place (STOIIP), MMSTB	353
Parameter of interest in model selection	STOIIP

Similarly, given that all the realizations have a similar profile to the reference model, it is not feasible to judge the accuracy and reliability of the realizations without an empirical analysis as detailed in Section 4.3.3.

4.3.3. Analysis of the Case Studies

Detailed SDP analysis and comparison of the divergence, signal-to-noise ratio, performance and uncertainties associated with each of the realizations of Field B are shown in Table 4.

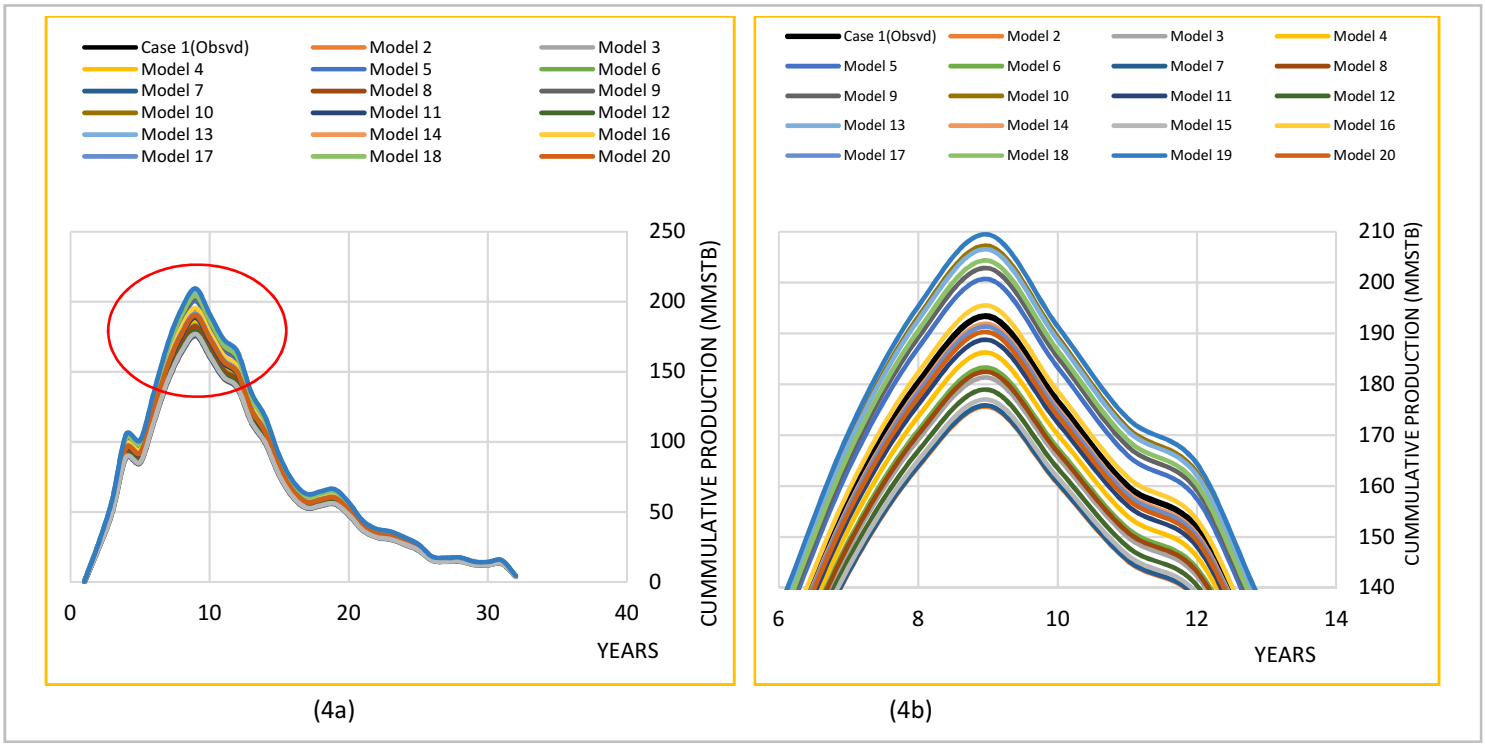

Figure 4: Production Profile of Field A (The figure on the right (4b) is the zoomed image of the circled section in Figure 4a)

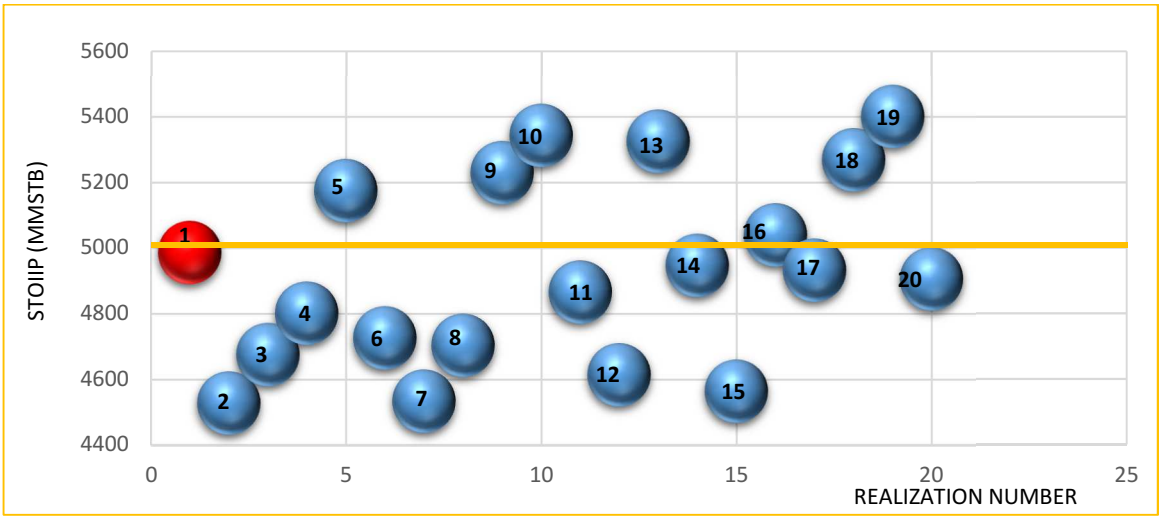

Figure 5: STOIIP distribution of Field A showing the Reference Model (1) and the Realizations

Table 3: Field A Model Ranking and Selection

FIELD-A												
	STOIIP (MMSTB)	**Dissimilarity measure (di)**	**Expected divergence (E)**	**Variance distribution**	**Signal (S)**	**Noise (N)**	**Signal-to-Noise ratio (SNR)**	**Reliability Factor (RF)**	**Mutual Divergence**	**Ranking by Divergence (Fitness)**	**Ranking by SDP (Reliability)**	**Rank Outcome**
Ref. Model	**4985.30**	0.00	0.00	0.05	0.95	0.05	17.60	0.98	2%	0	0	
Model 2	**4527.17**	3.08	0.16	0.04	0.96	0.04	21.56	1.20	-20%	19	19	-
Model 3	**4674.82**	3.05	0.15	0.05	0.95	0.05	20.16	1.12	-12%	12	12	-
Model 4	**4800.93**	3.02	0.15	0.05	0.95	0.05	19.06	1.06	-6%	7	6	-
Model 5	**5173.71**	2.95	0.15	0.06	0.94	0.06	16.27	0.90	10%	6	7	-
Model 6	**4725.45**	3.04	0.15	0.05	0.95	0.05	19.70	1.09	-9%	9	9	-
Model 7	**4533.02**	3.08	0.16	0.04	0.96	0.04	21.50	1.19	-19%	18	18	-
Model 8	**4703.38**	3.04	0.15	0.05	0.95	0.05	19.90	1.11	-11%	11	10	-
Model 9	**5229.16**	2.94	0.15	0.06	0.94	0.06	15.91	0.88	12%	8	8	-
Model 10	**5342.14**	2.91	0.15	0.06	0.94	0.06	15.20	0.84	16%	14	14	-
Model 11	**4865.90**	3.01	0.15	0.05	0.95	0.05	18.53	1.03	-3%	5	5	-
Model 12	**4613.60**	3.06	0.15	0.05	0.95	0.05	20.72	1.15	-15%	15	15	-
Model 13	**5324.45**	2.92	0.15	0.06	0.94	0.06	15.31	0.85	15%	13	13	-
Model 14	**4946.33**	2.99	0.15	0.05	0.95	0.05	17.90	0.99	1%	**1**	**1**	**2**
Model 15	**4563.05**	3.07	0.16	0.05	0.95	0.05	21.21	1.18	-18%	17	17	-
Model 16	**5039.46**	2.97	0.15	0.05	0.95	0.05	17.21	0.96	4%	**3**	**3**	**4**
Model 17	**4932.17**	2.99	0.15	0.05	0.95	0.05	18.01	1.00	0%	**2**	**2**	**1**
Model 18	**5267.59**	2.93	0.15	0.06	0.94	0.06	15.66	0.87	13%	10	11	-
Model 19	**5399.84**	2.90	0.15	0.06	0.94	0.06	14.86	0.83	17%	16	16	-
Model 20	**4905.47**	3.00	0.15	0.05	0.95	0.05	18.21	1.01	-1%	**4**	**4**	**2**

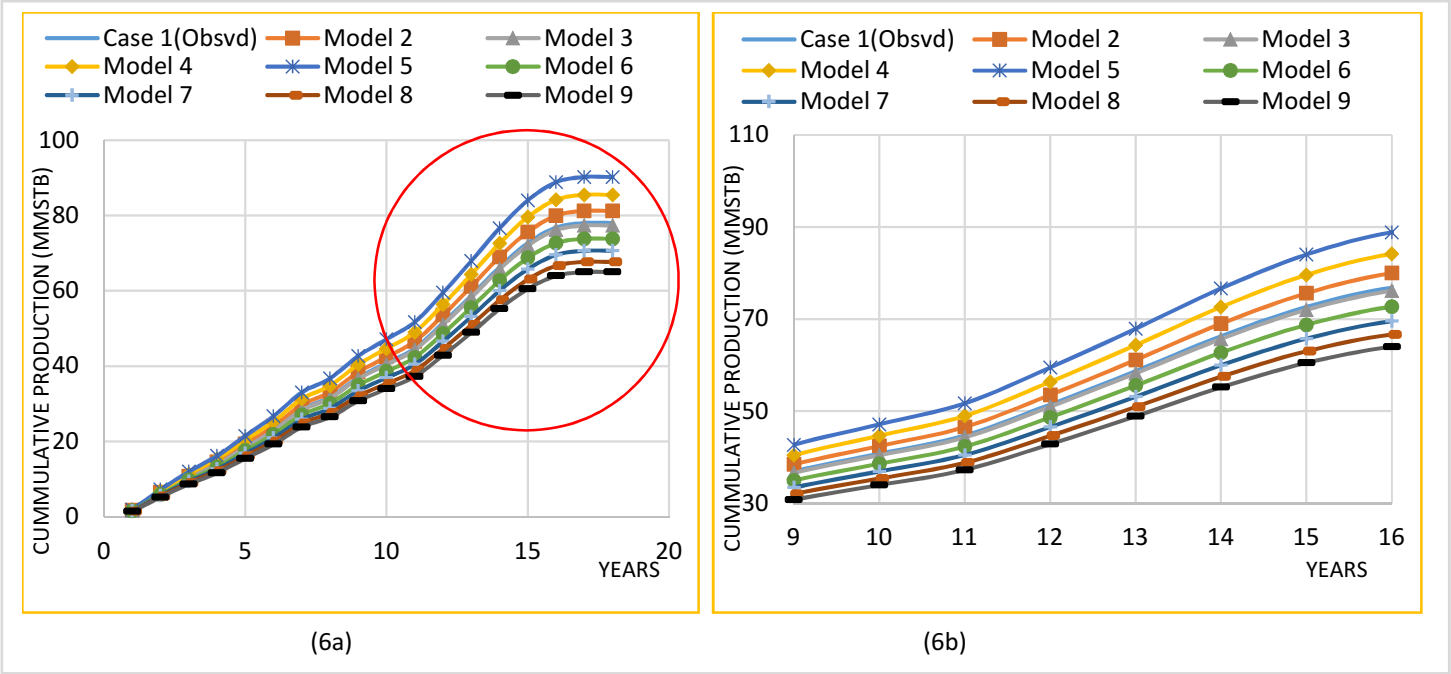

Figure 6: Production Profile of Field B (The figure on the right (6b) is the zoomed image of the circled section in Figure 6a)

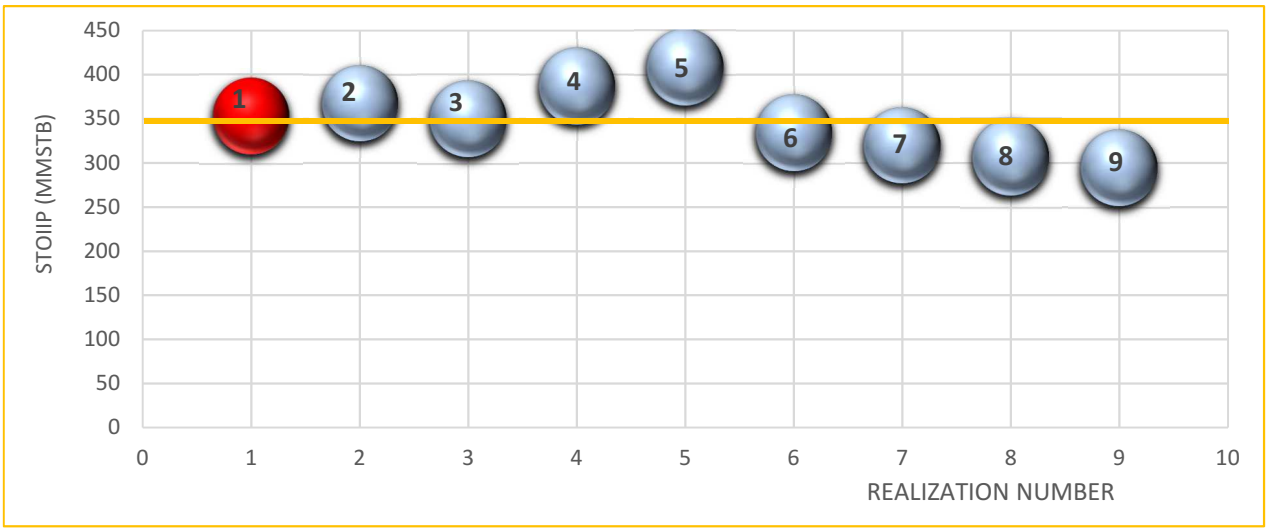

Figure 7: STOIIP distribution of Field B showing the Reference Model (1) and the Realizations

Table 4: Field B Model Ranking and Selection

FIELD-B												
	STOIIP (MMSTB)	Dissimilarity measure (di)	Expected divergence (E)	Variance distribution	Signal (S)	Noise (N)	Signal-to-Noise ratio (SNR)	Reliability Factor (RF)	Mutual Divergence	Ranking by Divergence (Fitness)	Ranking by SDP (Reliability)	Rank Outcome
Ref. Model	**353.04**	0.00	0.25	0.13	0.87	0.13	6.77	0.97	3%	0	0	
Model 2	**367.52**	2.14	0.24	0.14	0.86	0.14	6.17	0.88	12%	2	3	-
Model 3	**350.02**	2.19	0.25	0.13	0.87	0.13	6.90	0.99	1%	**1**	**1**	**1**
Model 4	**386.86**	2.09	0.24	0.15	0.85	0.15	5.47	0.78	22%	4	5	-
Model 5	**408.35**	2.03	0.23	0.17	0.83	0.17	4.80	0.69	31%	7	6	-
Model 6	**334.11**	2.23	0.25	0.12	0.88	0.12	7.67	1.10	-10%	3	2	**2**
Model 7	**319.58**	2.28	0.26	0.11	0.89	0.11	8.48	1.21	-21%	5	4	-
Model 8	**306.26**	2.32	0.26	0.10	0.90	0.10	9.32	1.33	-33%	6	7	-
Model 9	**294.01**	2.36	0.27	0.09	0.91	0.09	10.20	1.46	-46%	8	8	-

Table 5: Field A SDP Comparison with other Models

Field A	Actual Rank	SDP		KL Divergence		Mahalanobis Distance (Covariance Matrix)		Euclidean Distance	
MODELS		Value	Rank	Value	Rank	Value	Rank	Value	Rank
1	0	0.0221	0	0	0	0	0	0	0
2	19	-0.1977	19	0.00488	19	0.8064	19	73.3434	12
3	12	-0.1198	12	0.00325	12	0.35869	12	7.14384	4
4	6	-0.0588	6	0.00191	7	0.12316	7	23.2867	7
5	7	0.0960	7	0.00188	6	0.11935	6	111.847	14
6	9	-0.0947	9	0.00271	9	0.24854	9	12.61	6
7	18	-0.1945	18	0.00481	18	0.78493	18	0.01191	1
8	10	-0.1055	10	0.00294	11	0.29394	11	10.0519	5
9	8	0.1162	8	0.00242	8	0.19782	8	129.076	15
10	14	0.1555	14	0.0035	14	0.41461	14	166.684	18
11	5	-0.0293	5	0.00123	5	0.05096	5	34.7071	8
12	15	-0.1512	15	0.00392	15	0.5209	15	2.51334	3
13	13	0.1495	13	0.00333	13	0.37577	13	160.589	17
14	1	0.0057	1	0.0004	1	0.00534	1	51.4319	11
15	17	-0.1781	17	0.00448	17	0.67967	17	0.44266	2
16	3	0.0442	3	0.00055	3	0.01012	3	74.0122	13
17	2	-0.0003	2	0.00054	2	0.00996	2	48.2911	10
18	11	0.1299	11	0.00279	10	0.26315	10	141.507	16
19	16	0.1747	16	0.00404	16	0.55355	16	187.059	19
20	4	-0.0118	4	0.00082	4	0.0226	4	42.594	9
Match		**20/20**	**100%**	**16/20**	**80%**	**16/20**	**80%**	**1/20**	**5%**
Mis-Match		**0/20**	**0%**	**4/20**	**20%**	**4/20**	**20%**	**19/20**	**95%**

Table 6: Field B SDP Comparison with other Models

Field B	Actual Rank	SDP		KL Divergence		Mahalanobis Distance (Covariance Matrix))		Euclidean Distance	
MODELS		Value	Rank	Value	Rank	Value	Rank	Value	Rank
1	0	0.0334	0	0	0	0	0	0	0
2	3	0.1190	3	14.1858	2	0.10392	2	14.4747	2
3	1	0.0141	1	3.0391	1	0.00477	1	3.02606	1
4	5	0.2189	5	32.2945	4	0.53886	4	33.8176	5
5	6	0.3136	6	51.3824	7	1.36556	7	55.3098	7
6	2	-0.0960	2	19.4626	3	0.19563	3	18.9358	3
7	4	-0.2111	4	35.1559	5	0.63866	5	33.4622	4
8	7	-0.3314	7	50.1812	6	1.30235	6	46.7779	6
9	8	-0.4568	8	64.5930	8	2.16022	8	-59.028	8
Match		**9/9**	**100%**	**7/9**	**77.70%**	**7/9**	**77.70%**	**3/9**	**33.33%**
Mis-Match		**9/9**	**0%**	**2/9**	**22.22%**	**2/9**	**22.22%**	**6/9**	**66.66%**

Considering that the realizations were equiprobably generated with some probability distribution function, Tables 7 and 8 specify the mean, standard deviation and variance of the distribution for Field A and Field B respectively.

Table 7: Field-A Probability Distribution

FIELD-A			
	Mean	Std Dev	Variance
Case 1(Obsvd)	71.47	58.91	3470.05
Model 2	64.90	53.49	2861.59
Model 3	67.02	55.24	3051.28
Model 4	68.83	56.73	3218.12
Model 5	74.17	61.13	3737.29
Model 6	67.75	55.84	3117.74
Model 7	64.99	53.56	2868.99
Model 8	67.43	55.58	3088.68
Model 9	74.97	61.79	3817.84
Model 10	76.59	63.12	3984.59
Model 11	69.76	57.50	3305.82
Model 12	66.14	54.52	2971.89
Model 13	76.33	62.91	3958.24
Model 14	70.91	58.45	3416.01
Model 15	65.42	53.92	2907.12
Model 16	72.25	59.55	3545.85
Model 17	70.71	58.28	3396.48
Model 18	75.52	62.24	3874.16
Model 19	77.41	63.81	4071.12
Model 20	70.33	57.96	3359.81

Table 8: Field-B Probability Distribution

FIELD-B			
	Mean	Std Dev	Variance
Case 1(Obsvd)	41.02	26.14	683.34
Model 2	42.70	27.21	740.53
Model 3	40.66	25.92	671.68
Model 4	44.94	28.64	820.53
Model 5	47.44	30.24	914.23
Model 6	38.82	24.74	612.01
Model 7	37.13	23.66	559.94
Model 8	35.58	22.68	514.25
Model 9	34.16	21.77	473.94

4.4. General Discussions

4.4.1. Result Comparison

Comparison of the results of the realizations in Field A with its reference model shows that despite realization 14 fitting the reference model closely as shown in Figure 5, it was not the most reliable model for future forecasts. The reason for this is simple: the geostatistical gradient-based and distance-based model ranking methods are unable to capture the fudge factor that is associated with individual realizations. These methods did not consider the possibility that there are other factors that could make any model or realization agree with a reference model; this is an anomaly and hence results in unreliable predictions of the real system. The same reason applies to Field B as shown in Figure 7 with realizations 3 ranking as the best fit; and realization 6 ranking higher than realization 2 despite the similarity of their fitness to the reference model. It is therefore important to mention here that one of the reasons for the anomaly of erroneous model fitness is the non-uniqueness of a model's parameter set. This is a mathematical problem which states that there is no unique set of parameters that could be the only parameter set to fit any model, but that there will always exist many other parameter sets that could fit the same model under different conditions. Therefore, different combinations of parameters from other models and realizations will be able to match a reference model if they are able to arrive at similar gradients. The overall comparison of the results of the SDP method to other popular and widely-accepted distance-based model ranking methods (Tables 5 and 6) shows that the SDP approach makes better and more reliable ranking and predictions.

The *Actual rank* used in the analysis in Tables 3 and 4 is based on the manual testing and validation of each realization with respect to its accuracy in the prediction and forecast of the real-life field data. Given that the reference model in this study represents the actual data; all ranking was done with reference to it. In Tables 3 to 6, it is shown that all the ranks for the reference model (being the actual data) is zero, whereas the realizations of the reference model are ranked from number 1 upwards, with the strongest model having the rank of 1.

Regarding the results obtained from the two reservoir examples in this study, the reliability factors (RF) obtained for each example shows that the choice of nineteen realizations for Field A is more accurate than the choice of nine realizations for Field B. For Field A, an optimal model (Realization 17) with an RF of 1 and a Mutual Divergence of 0% was achieved; and the SDP results also revealed that the Field A reference model contained some inherent fudge of about 2%. On the other hand for Field B, Realization 3 with an RF of 0.99 and a Mutual Divergence of 1% from the reference model emerged as the closest and optimal model; and the SDP results reveal that the Field B reference model had an inherent fudge of about 3%.

This shows that there is a definite cut-off for the number of realizations required to optimally capture the full range of uncertainty in the reservoir model, in order to avoid either underfitting or overfitting the model.

5.0. Conclusion

Several methods have been applied in the ranking and selection of reservoir models, with each approach having its own merits and flaws; yet the issue of selecting reliable predictive models or realizations, which gives least uncertainty, has lingered. In addition, quantifying model uncertainties and characterizing them has remained a challenge. This paper has introduced the SDP approach to resolve this issue. SDP ranks the realizations using a dissimilarity factor, Signal-to-Noise ratio (SNR) and a Reliability Factor (RF) which is a measure of the accuracy of each realization. The model with a reliability factor of one is chosen as the most reliable, while

models with reliability factors that diverge from the value of one are more unreliable in making predictions and forecasts. Results from the two case examples treated in this paper have shown that SDP can rank and select the most suitable model from an ensemble of multiple realizations. Unlike other model ranking and selection methods, the uniqueness of the SDP approach is based on its ability to identify any inherent noise in presumably good and fitted models, in addition to quantifying the model uncertainties as shown in this paper.

In addition, with regards to the basis for deciding on the optimum number of realizations that can fully capture the entire uncertainty envelope in any model space, this paper has clearly demonstrated that if more or less number of realizations than necessary are used, the error margin would increase rather than decrease; thereby adding more uncertainty into the output of the model being analysed.

Finally, with the SDP approach, the actual impact of the mismatch uncertainty of a model, and other associated uncertainties can be characterized and quantified.

Acknowledgement

We are particularly grateful to the Petroleum Technology Development Fund (PTDF) Nigeria, for providing the funding for this research.

References

Ates, H., Bahar, A., El-Abd, S., Charfeddine, M., Kelkar, M., and Datta-Gupta, A. "Ranking and Upscaling of Geostatistical Reservoir Models Using Streamline Simulation: A Field Case Study." *Middle East Oil Show, Bahrain,* Society of Petroleum Engineers, 9-12 June, 2003.

Ballin, P.R., Journel, A.G. and Aziz, K. (1992): "Prediction of uncertainty in reservoir performance forecast", *Journal of Canadian Petroleum Technology,* April 1992*,* vol. 31, no. 04, pp. 52-62.

Begg, S.H. and Welsh, M.B. "Uncertainty vs. Variability: What's the Difference and Why is it Important?", *SPE Hydrocarbon Economics and Evaluation Symposium, Houston, Texas,* Society of Petroleum Engineers, 19-20 May, 2014.

Cannon, S. (2018), *Reservoir Modelling: A practical Guide*, Wiley Blackwell, New Jersey.

Dehdari, V. and Deutsch, C.V., " Ranking Reservoir Realizations Using SAGD Proxy and Gradual Deformation Method", *Centre for Computational Geostatistics Report*, 2012, 14, 204, pp. 1-14, University of Alberta, Canada.

Deutsch, C. V. and Journel., A. G., (1998): *GSLIB Geostatistical Software Library and User's Guide,* 2nd Ed., New York: Oxford University Press, 119–122.

Deutsch, C.V. and Srinivasan, S. "Improved Reservoir Management Through Ranking Stochastic Reservoir Models", *SPE/DOE Improved Oil Recovery Symposium, Tulsa, Oklahoma,* Society of Petroleum Engineers, 21-24 April, 1996.

Fanchi, J.R. (2006), *Principles of Applied Reservoir Simulation,* 3rd edn, Elsevier, Oxford.

Fei, J., Jeffrey, M., Yarus, M. and Chambers, R. 2016, "Apply Two-Way Cluster Analysis to Select Candidate Reservoir Models from Multiple Realizations.", *SPE/IAEE Hydrocarbon Economics and Evaluation Symposium, Houston, Texas, USA,* Society of Petroleum Engineers, 17-18 May.

Fenwick, D. and Batycky, R., "Using Metric Space Methods to Analyse Reservoir Uncertainty", *2011 Gussow Geoscience Conference, Banff, Alberta,* Canadian Society of Petroleum Geologists (CSPG).

Guerra, N. Y. and Narayanasamy, R., "Well Location Selection from Multiple Realisations of a Geomodel Using Productivity Potential Maps - A Heuristic Technique." *International Oil Conference and Exhibition in Mexico, Cancun, Mexico,* Society of Petroleum Engineers, 31 August-2 September, 2006.

Hegstad, B. K. and Saetrom, J., "Using Multiple Realizations from an Integrated Uncertainty Analysis to Make More Robust Decisions in Field Development." *Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE,* Society of Petroleum Engineers, 10-13 November, 2014.

Idrobo, E. A., Choudhary, M. K., and Datta-Gupta, A. "Swept Volume Calculations and Ranking of Geostatistical Reservoir Models Using Streamline Simulation", *SPE/AAPG Western Regional Meeting, Long Beach, California,* Society of Petroleum Engineers, 19-22 June, 2000.

Li, S. and Deutsch, C.V., "A Petrel Plugin for Ranking Realizations", *Centre for Computational Geostatistics Report*, 2008, 10, 407, pp. 1-8, University of Alberta, Canada.

Li, S., Deutsch, C.V. and Si, J. "Ranking Geostatistical Reservoir Models with Modified Connected Hydrocarbon Volume", *Ninth International Geostatistics Congress, Oslo, Norway,* June 11 – 15, 2012.

McLennan, J.A. and Deutsch, C.V., "Ranking Geostatistical Realizations by Measures of Connectivity", *SPE International Thermal Operations and Heavy Oil Symposium, Calgary, Alberta, Canada,* Society of Petroleum Engineers, 1-3 November 2005.

McLennan, J.A. and Deutsch, C.V., "Ranking Geostatistical Realizations for Flow Simulation", *Centre for Computational Geostatistics Report,* 2005, 7, 114, pp. 1-22, University of Alberta, Canada.

Mtchedlishvili, G., Voigt, H. and Haefner, F., "Model Selection and Inference for Reservoir History Matching", *9th European Conference on the Mathematics of Oil Recovery, Cannes, France,* European Association of Geoscientists and Engineers, 30 August - 2 September, 2004.

Pratt, H. R., Robey, E. H., Wojewodka, R. A., and Mercer, J. C. "Coalbed Methane Geostatistical Analysis: A Method for Ranking Well Productivity", *SPE Unconventional Gas Technology Symposium, Louisville, Kentucky,* Society of Petroleum Engineers, 18-21 May, 1986.

Rahim, S., Li, Z. and Trivedi, J. (2015), "Reservoir Geological Uncertainty Reduction: An Optimization-Based Method Using Multiple Static Measures", *Mathematical Geosciences,* vol. 47, no. 4, pp. 373-396.

Saad, N., Maroongroge, V. and Kalkomey,C., "Ranking Geostatistical Models Using Tracer Production Data", *European 3-D Reservoir Modelling Conference, Stavanger, Norway,* Society of Petroleum Engineers, , 16-17 April, 1996.

Sahni, I., and Horne, R. N. "Generating Multiple History-Matched Reservoir Model Realizations Using Wavelets", *SPE Annual Technical Conference and Exhibition, Houston, Texas,* Society of Petroleum Engineers, 26-29 September, 2004.

Scheidt, C. and Caers, J. (2010), "Bootstrap confidence intervals for reservoir model selection techniques", *Computational Geosciences,* vol. 14, no. 02, pp. 369-382.

Sharifi, M., Kelkar, M., Bahar, A., and Slettebo, T. (2014), "Dynamic Ranking of Multiple Realizations By Use of the Fast-Marching Method", *SPE Journal,* December 2014, pp. 1069 - 1082.

Shirangi, M.G. (2014), "History matching production data and uncertainty assessment with an efficient TVSD parameterisation algorithm", *Journal of Petroleum Science and Engineering,* vol. 113, pp. 54-71.

Shirangi, M.G. and Durlofsky, L.J. (2016), "A generalised method to select representative models for decision making and optimization under uncertainty", *Computers and Geosciences,* vol. 96, pp. 109-123.

Shirangi, M.G. and Emerick, A.A. (2016), "An improved TVSD-based Levenberg-Marquardt algorithm for history matching and comparison with Gauss-Newton", *Journal of Petroleum Science and Engineering,* vol. 143, pp. 258-271.

Wang, Y. and Kovscek, A. "A Streamline Approach for Ranking Reservoir Models that Incorporates Production History", *SPE Annual Technical Conference and Exhibition, San Antonio, Texas,* Society of Petroleum Engineers, 29 September - 2 October, 2002.

Yadav, S., Bryant, S. L., and Srinivasan, S., "Ranking of Geostatistical Reservoir Models and Uncertainty Assessment Using Real-Time Pressure Data", *SPE Western Regional/AAPG Pacific Section/GSA Cordilleran Section Joint Meeting, Anchorage, Alaska, USA,* Society of Petroleum Engineers, 8-10 May, 2006.